DOCUMENT RESUME

ED 473 962 CS 511 919

AUTHOR	Ogle, Laurence T.; Sen, Anindita; Pahlke, Erin; Jocelyn, Leslie; Kastberg, David; Roey, Stephen; Williams, Trevor
TITLE	International Comparisons in Fourth-Grade Reading Literacy: Findings from the Progress in International Reading Literacy Study (PIRLS) of 2001.
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.; Education Statistics Services Inst., Washington, DC.; Westat, Inc., Rockville, MD.
REPORT NO	NCES-2003-073
PUB DATE	2003-04-00
NOTE	41p.
AVAILABLE FROM	National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, 1990 K St., NW, Washington, DC 20006. Tel: 800-283-6237 (Toll Free). For full text: http://www.nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2003073.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Comparative Analysis; Educational Indicators; Foreign Countries; *Grade 4; Intermediate Grades; International Programs; *Literacy; Low Income; Racial Differences; *Reading Achievement; Reading Attitudes; Reading Tests; Recreational Reading; Sex Differences
IDENTIFIERS	International Assn Evaluation Educ Achievement; Reading Behavior

ABSTRACT

The Progress in International Reading Literacy Study (PIRLS) is an assessment of reading comprehension of nine-year-olds in 35 countries and was conducted by the International Association for the Evaluation of Educational Achievement (IEA). This report compares findings from the United States with other countries and takes a closer look at performance within the United States. Results indicated that: (1) United States fourth graders outperformed their counterparts in 23 of the 34 other countries, although they scored lower than students in England, the Netherlands, and Sweden; (2) the percentage of U.S. fourth graders reaching each of four benchmarks is higher than the international averages; (3) fourth-grade girls scored higher than fourth-grade boys in the combined reading literacy scale on average in every participating country; (4) with the exception of Black fourth graders, each racial/ethnic group in the United States scored higher than the international average on the combined reading literacy scale, as well as on the two reading subscales; (5) fourth graders in U.S. public elementary schools with the highest poverty levels scored lower on the combined reading literacy scale compared to their counterparts in schools with lower poverty levels; (6) 95% of U.S. fourth-grade students attend schools with a curricular emphasis on reading, which is greater than the international average of 78%; (7) 35% of U.S. fourth-grade students reported reading for fun every day or almost every day, which is smaller than the international average of 40%; (8) 32% of U.S. fourth-grade students reported they never or almost never read for fun outside of school, a significantly higher percentage that the international average of 18%; and (9) U.S. fourth graders

Reproductions supplied by EDRS are the best that can be made
from the original document.

who reported reading literary fiction outside of school at least once or twice a month had higher scores on the combined reading literacy scale than those who never or almost never did so. Contains sample items from PIRLS 2001 and a brief description of the IEA International Reading Literacy Study of 1991. Appendixes contain technical notes; a brief analysis of the similarities and differences between PIRLS and the upcoming National Assessment of Educational Progress results of testing of U.S. fourth graders; and supplemental tables of data. (Contains 15 figures and 10 exhibits.) (RS)

Reproductions supplied by EDRS are the best that can be made
from the original document.

NCES
National Center for Education Statistics

PIRLS
Progress in International Reading Literacy Study

International Comparisons in Fourth-Grade Reading Literacy:

Findings from the Progress in International Reading Literacy Study (PIRLS) of 2001

U.S. Department of Education
Institute of Education Sciences
NCES 2003–073

ED 473 962

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS 511 919

BEST COPY AVAILABLE

NCES
National Center for Education Statistics

PIRLS Progress in International Reading Literacy Study

U.S. Department of Education
Institute of Education Sciences
NCES 2003–073

International Comparisons in Fourth-Grade Reading Literacy:

Findings from the Progress in International Reading Literacy Study (PIRLS) of 2001

April 2003

Laurence T. Ogle
National Center for Education Statistics

Anindita Sen
Erin Pahlke
Education Statistics Services Institute

Leslie Jocelyn
David Kastberg
Stephen Roey
Trevor Williams
Westat

U.S. Department of Education
Rod Paige
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Val Plisko
Associate Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006

April 2003

The NCES World Wide Web Home Page is: **http://nces.ed.gov**
The NCES World Wide Electronic Catalog is: **http://nces.ed.gov/pubsearch**

Suggested Citation
Ogle, L., Sen, A., Pahlke, E., Jocelyn, L., Kastberg, D., Roey, S., and Williams, T. (2003). *International Comparisons in Fourth-Grade Reading Literacy: Findings from the Progress in International Reading Literacy Study (PIRLS) of 2001* (NCES 2003–073). U.S. Department of Education, NCES. Washington, DC: U.S. Government Printing Office.

For ordering information on this report, write:

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

or call toll free 1-877-4ED-PUBS or go to the Internet: **http://www.edpubs.org**

Content Contact:
Laurence Ogle: 202-502-7426
Fax: 202-502-7466
E-mail: Laurence.Ogle@ed.gov

Acknowledgments
Many people contributed to making this report possible, and the authors wish to thank all those who have assisted with various aspects of the report, including data analysis, review, and design.

We would like to thank Shelley Burns, Peggy Carr, Kathy Chandler, Patrick Gonzales, Elvira Germino Hausken, Mariann Lemke, Karen O'Conor, Val Plisko, and Marilyn Seastrom of NCES for their helpful technical and editorial feedback. Invited reviewers who gave of their time and expertise include Bonnie Armbruster of the University of Illinois, Jean Osborn, education consultant, and Anne Sweet of the Institute of Education Sciences, U.S. Department of Education.

Eugenio Gonzalez and Ann Kennedy of the International Study Center at Boston College also provided valuable information about the PIRLS data.

Finally, Tracey Summerall of the Westat Graphic Arts Department designed the report cover and graphics throughout the report.

BEST COPY AVAILABLE

Table of Contents

Acknowledgments iii

List of Figures vii

List of Exhibits viii

PIRLS 2001 1

PIRLS 2001 in Brief 2

Reading Literacy Defined 3

Reading Literacy 4

U.S. Student Performance on PIRLS 2001 4

U.S. Student Performance on the Combined Reading Literacy Scale . . . 4

U.S. Student Performance on Subscales 4

Distribution of Average Combined Reading Literacy Scores 6

Reading Literacy by Benchmarks 8

How Different Groups Perform 10

Achievement by Sex 10

U.S. Achievement by Race/Ethnicity 13

U.S. Achievement by Control of School 14

U.S. Achievement by Poverty Level in Public Schools 14

Reading and Instruction in the Classroom 16

Reading Curriculum and Instructional Time 16

Teacher Preparation and Experience 17

Reading Outside of School 18

Reading Outside of School for Enjoyment 18

Choice of Activities Outside of School 19

Sample Items from PIRLS 2001 20

IEA International Reading Literacy Study of 1991 24

Reading Performance Over Time 24

Performance on the IEA International Reading Literacy Study of 1991 25

Works Cited 26

Appendix A, Technical Notes 27

Appendix B, Examining the Similarities and Differences Between PIRLS and NAEP 30

List of Figures

Figure 1. Participating countries in PIRLS 2001 2

Figure 2. Percentage of PIRLS assessment items devoted to reading purposes and processes 3

Figure 3. Fourth-graders' average scores for the combined reading literacy scale, literary subscale, and informational subscale, by country: 2001 5

Figure 4. Distribution of average combined reading literacy scale scores of fourth-graders by percentiles, by country: 2001 7

Figure 5. Fourth-graders' reading skills and strategies, and cut point scores, by benchmark points for the combined reading literacy scale: 2001 8

Figure 6. Percentage of fourth-grade students who reach the PIRLS 2001 achievement benchmarks: 2001 9

Figure 7. Difference in average scores between boys and girls for the combined reading literacy scale of fourth-graders, by country: 2001 11

Figure 8. Fourth-graders' average scores on the combined reading literacy scale, literary subscale, and informational subscale for boys and girls: 2001 12

Figure 9. U.S. fourth-graders' average scores for the combined reading literacy scale, literary subscale, and informational subscale, by race/ethnicity: 2001 13

Figure 10. U.S. public school fourth-graders' average scores for the combined reading literacy scale, by the percentage of school enrollment eligible for free or reduced-price lunch: 2001 15

Figure 11. Percentage of fourth-graders by average number of hours of reading instruction each week: 2001 16

Figure 12. Average number of years fourth-grade students' teachers have taught all grades and fourth grade: 2001 17

Figure 13. Percentage of fourth-grade students who read for fun outside of school, by frequency of reading: 2001 18

Figure 14. Participating countries in the IEA International Reading Literacy Study of 1991: 2001 24

Figure 15. Average scores for the IEA International Reading Literacy Study of 1991 combined reading literacy scale and differences in average achievement scores for fourth-graders, by country: 1991 and 2001 25

List of Exhibits

Exhibit 1. Questions PIRLS 2001 can answer 1

Exhibit 2. How PIRLS 2001 was constructed and administered 2

Exhibit 3. Aspects of reading literacy assessed in PIRLS 2001 3

Exhibit 4. Example of item at the top 10 percent PIRLS 2001 international benchmark 21

Exhibit 5. Example of item at the top 10 percent PIRLS 2001 international benchmark 21

Exhibit 6. Example of item at the upper quarter PIRLS 2001 international benchmark 22

Exhibit 7. Example of item at the median PIRLS 2001 international benchmark 22

Exhibit 8. Example of item at the lower quarter PIRLS 2001 international benchmark 23

Exhibit 9. Example of item at the lower quarter PIRLS 2001 international benchmark 23

Exhibit 10. Interpreting PIRLS 2001 in light of the IEA International Reading Literacy Study of 1991 24

PIRLS 2001

The Progress in International Reading Literacy Study (PIRLS) is an assessment of reading comprehension conducted by the International Association for the Evaluation of Educational Achievement (IEA). Thirty-five countries assessed the reading literacy of their students in the upper of the two grades with the most 9-year-olds (fourth grade in most countries, including the United States). PIRLS 2001 provides comparative information on the reading literacy of these fourth-graders and also examines factors that may be associated with the acquisition of reading literacy in young children (see exhibit 1).

As the sponsor for PIRLS 2001 in the United States, the National Center for Education Statistics (NCES) is reporting findings from the study that compare the United States with other countries and that take a closer look at performance within the United States. The full report on the international study is available at **www.pirls.org**. Also available at this site is the *PIRLS Technical Report*, which examines specific technical issues related to the assessment. Supporting data for the tables and analyses in this report are available at **www.nces.ed.gov/surveys/pirls**.

Exhibit 1. Questions PIRLS 2001 can answer

PIRLS 2001 is a carefully constructed reading assessment, consisting of a test of reading literacy and questionnaires to collect information about fourth-grade children's reading literacy performance.

PIRLS 2001 will help educators and policymakers by answering questions such as the following:

- How well do fourth-grade students read?
- How do students in one country compare with students in another country?
- Do fourth-grade students value and enjoy reading?
- Internationally, how do the reading habits and attitudes of students vary?

BEST COPY AVAILABLE

PIRLS 2001 in Brief

PIRLS 2001 follows by 10 years a prior IEA study of reading literacy called the IEA International Reading Literacy Study of 1991. Over the 10 years between these studies, progress has been made in the ways in which students are assessed and in the construction of the assessment instruments themselves. There has also been a shift in the design of the assessments. Thus, while PIRLS 2001 can trace its evolution from the earlier IRLS, it is nevertheless a different study.

PIRLS 2001 is the first in a planned 5-year cycle of international trend studies in reading literacy by the IEA. PIRLS is designed to assist participating countries in monitoring the reading literacy of their fourth-grade populations in comparison to other countries.

Thirty-five countries participated in PIRLS 2001 (see figure 1). The survey assessed the reading literacy of children in the upper of the two grades with the most 9-year-olds—fourth grade in most countries, including the United States (exhibit 2). (See appendix A for more information on the sampling in PIRLS 2001.)

Data were collected in the final months of the 2000–01 school year. In the United States, data were collected in the spring of 2001 from both public and private schools.

Figure 1. Participating countries in PIRLS 2001

Argentina	Hungary	New Zealand
Belize	Iceland	Norway
Bulgaria	Iran	Romania
Canada (O,Q)[1]	Israel	Russian Federation
Colombia	Italy	Scotland
Cyprus	Kuwait	Singapore
Czech Republic	Latvia	Slovak Republic
England	Lithuania	Slovenia
France	Macedonia	Sweden
Germany	Moldova	Turkey
Greece	Morocco	**United States**
Hong Kong, SAR[2]	Netherlands	

[1] Canada is represented by the provinces of Ontario and Quebec (O, Q) only.

[2] Hong Kong, SAR is a Special Administrative Region (SAR) of the People's Republic of China.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

Exhibit 2. How PIRLS 2001 was constructed and administered

A group of distinguished international reading scholars, the Reading Development Group, was formed to construct the *PIRLS 2001 Framework* and endorse the final reading assessment. Each country followed internationally prescribed procedures to ensure valid translations and representative samples of students. Quality Control Monitors were then appointed in each country to monitor the testing sessions at the schools to ensure that the high standards of the PIRLS 2001 data collection process were met.

Reading literacy achievement was measured by using a selection of four literary passages drawn from children's storybooks and four informational texts.

Submitted and reviewed by the PIRLS 2001 countries, the literary passages included realistic stories and traditional tales. The informational texts included chronological and nonchronological articles, a biographical article, and an informational leaflet.

For more information on the test construction, see appendix A. More details are also available in the *PIRLS 2001 Technical Report* at **www.pirls.org.**

BEST COPY AVAILABLE

Reading Literacy Defined

PIRLS 2001 measures reading abilities at a time in students' schooling when most have learned how to read and are now using reading to learn.

PIRLS 2001 defines reading literacy as the following:

> *The ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers, and for enjoyment* (Campbell et al., 2001, p.3).

In PIRLS 2001, *purposes of reading* (see exhibit 3) refers to the two types of reading that account for most of the reading young students do, both in and out of school: (1) reading for literary experience, and (2) reading to acquire and use information. In the assessment, narrative fiction is used to assess students' ability to read for literary experience, while a variety of informational texts are used to assess students' ability to acquire and use information while reading. The PIRLS 2001 assessment contains an equal proportion of text assessing each purpose.

Processes of comprehension refer to ways in which readers construct meaning from the text. Readers focus on and retrieve specific ideas, make inferences, interpret and integrate ideas and information, and examine or evaluate text features.

Purposes of reading and processes of comprehension are the foundations of the assessment. Figure 2 portrays the interaction of the two: each process is assessed within each purpose of reading.

Exhibit 3. Aspects of reading literacy assessed in PIRLS 2001

1. Purposes of reading
2. Processes of comprehension
3. Reading behavior and attitudes

The first two aspects of reading literacy form the basis of the written test of reading comprehension. The student background questionnaire addresses the third aspect.

Figure 2. Percentage of PIRLS assessment items devoted to reading purposes and processes

	Purpose of reading (percent)		
Process of comprehension	Literary items	Informational items	Total
Total	49	50	100
Focus on and retrieve explicitly stated information	9	13	22
Make straightforward inferences	14	9	23
Interpret and integrate ideas and information	20	20	40
Examine and evaluate content, language, and textual elements	6	8	14

NOTE: Detail may not sum to totals due to rounding.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study, 2001.

Reading Literacy

U.S. Student Performance on PIRLS 2001

PIRLS 2001 scores are reported on a scale of 0 to 1000, with an international average of 500 and a standard deviation of 100.[1] For the 35 countries that participated in PIRLS 2001, figure 3 presents the average scores for the three scales: the combined reading literacy scale and its two components, the literary and informational subscales.[2]

U.S. Student Performance on the Combined Reading Literacy Scale

- U.S. fourth-grade students perform significantly better than the international average of 500 on the combined reading literacy scale (figure 3).[3]
- U.S. fourth-graders outperform their counterparts in 23 of the 34 other countries participating in PIRLS 2001, although they score lower than students in England, the Netherlands, and Sweden. No detectable differences in scores are found between U.S. students and their counterparts in eight of the remaining PIRLS 2001 countries.

U.S. Student Performance on Subscales

- U.S. fourth-grade students perform better than the international averages on both of the reading subscales.
- Sweden outscores the United States on the literary subscale, and five countries—Bulgaria, England, Latvia, the Netherlands, and Sweden—outperform the United States on the informational subscale.
- U.S. fourth-graders outscore students in 26 countries on the literary subscale and outperform their counterparts in 17 countries on the informational subscale.

[1] The international average is the mean of all countries participating in the study calculated so that all participating countries have the same contribution to the average. The PIRLS 2001 scale average for each scale (the combined reading literacy scale and the literary and informational subscales) across countries was set to 500 and the standard deviation to 100. Since the countries varied in size, each country was weighted to contribute equally to the mean and standard deviation of the scales. The average and standard deviation of each of the scale scores are arbitrary and do not affect scale interpretation. The standard deviation is the statistical measure of the extent to which values are spread around the average.

[2] Average scores for each country are based on a sample of students, rather than all students, and are estimates of the population value of all 9-year-olds in each country. These estimates have a known degree of sampling error, the standard error, and an unknown degree of nonsampling error. The true average for any country lies within a range of approximately two times the standard error above and below the estimated score (also known as the confidence interval). The combined literacy scale is based on the distribution of scores on all the test items, while the subscales are based on only the items that belong to each subscale. Hence, the combined reading literacy score is not the statistical average of the scores of the two subscales.

[3] All differences discussed in this report are statistically significant. No statistical adjustments (such as Bonferonni) are made while carrying out multiple comparisons between the United States and other countries. The *t*-tests used (in this report) do not adjust for the correlation between the U.S. average and the international average in order to be consistent with the comparisons carried out for the international report.

Figure 3. Fourth-graders' average scores for the combined reading literacy scale, literary subscale, and informational subscale, by country: 2001

Country	Average combined reading literacy score	Country	Average literary subscale score	Country	Average informational subscale score
Sweden	561	Sweden	559	Sweden	559
Netherlands[1]	554	England[1,4a]	559	Netherlands[1]	553
England[1,4a]	553	Netherlands[1]	552	Bulgaria	551
Bulgaria	550	**United States[1]**	**550**	Latvia	547
Latvia	545	Bulgaria	550	England[1,4a]	546
Canada (O, Q)[3,5]	544	Hungary	548	Canada (O, Q)[3,5]	541
Lithuania[3]	543	Lithuania[3]	546	Lithuania[3]	540
Hungary	543	Canada (O, Q)[3,5]	545	Germany	538
United States[1]	**542**	Italy	543	Hungary	537
Italy	541	Latvia	537	Hong Kong, SAR[6]	537
Germany	539	Germany	537	Czech Republic	536
Czech Republic	537	Czech Republic	535	Italy	536
New Zealand	529	New Zealand	531	**United States[1]**	**533**
Scotland[1]	528	Scotland[1]	529	France	533
Singapore	528	Singapore	528	Russian Federation[4a]	531
Russian Federation[4a]	528	Greece[1]	528	Singapore	527
Hong Kong, SAR[6]	528	Russian Federation[4a]	523	Scotland[1]	527
France	525	Iceland	520	New Zealand	525
Greece[1]	524	France	518	Slovak Republic	522
Slovak Republic	518	Hong Kong, SAR[6]	518	Greece[1]	521
Iceland	512	Slovak Republic	512	Romania	512
Romania	512	Romania	512	Israel[4a]	507
Israel[1]	509	Israel[1]	510	Moldova	505
Slovenia	502	Norway	506	Iceland	504
Norway	499	Slovenia	499	Slovenia	503
Cyprus	494	Cyprus	498	Norway	492
Moldova	492	Moldova	480	Cyprus	490
Turkey	449	Turkey	448	Turkey	452
Macedonia	442	Macedonia	441	Macedonia	445
Colombia	422	Colombia	425	Colombia	424
Argentina	420	Iran	421	Argentina	422
Iran	414	Argentina	419	Iran	408
Kuwait	396	Kuwait	394	Kuwait	403
Morocco[2]	350	Morocco[2]	347	Morocco[2]	358
Belize	327	Belize	330	Belize	332
International Average	500	International Average	500	International Average	500

Average is significantly higher than the U.S. average

Average is not significantly different from the U.S. average

Average is significantly lower than the U.S. average

[1] Met guidelines for sample participation rates after replacement schools were included.

[2] Nearly satisfying guidelines for sample participation rates after replacement schools were included.

[3] National Desired Population does not cover all of International Desired Population because coverage falls below 65 percent.

[4a] National Defined Population covers less than 95 percent of National Desired Population.

[4b] National Defined Population covers less than 80 percent of National Desired Population.

[5] Canada is represented by the provinces of Ontario and Quebec (O, Q) only.

[6] Hong Kong, SAR is a Special Administrative Region (SAR) of the People's Republic of China.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

BEST COPY AVAILABLE

Distribution of Average Combined Reading Literacy Scores

The average scores for reading literacy describe how a country performs overall compared to other nations, but they provide no information about the way scores are distributed within the countries. One country with an average score similar to another could have large numbers of high- and low-scoring students, while the other country could have large numbers of students performing at about the average score. Figure 4 details how scores were distributed across countries.

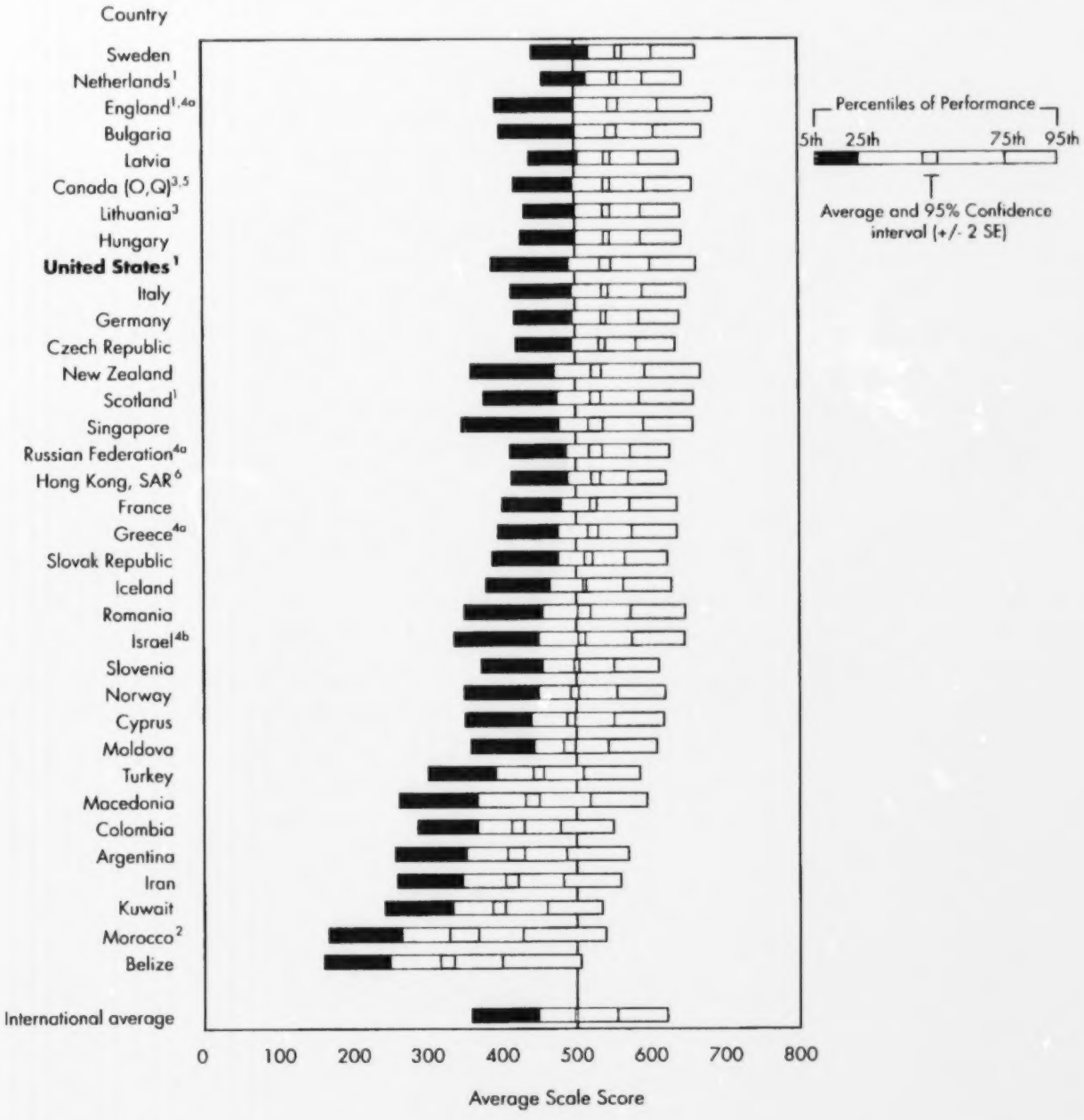
- In the United States, the 5th percentile score forcombined reading literacy is 389. Ninety-five percent of U.S. students score above 389; in the same way, 5 percent of students score above 663, the 95th percentile score. This means that the top 5 percent of U.S. students score at least 274 points higher than the bottom 5 percent (figure 4).

Looking at the length of the bars in figure 4 gives a sense of how large the differences are between a country's highest and lowest performing students, but it does not describe how many students are high or low performing. As with average scores, because of the statistical techniques used to sample students, it is not accurate to rank countries' scoring variation based simply on the length of the bars shown in figure 4. Standard deviations of the combined reading literacy average scores gives a mathematical way to tell how greatly scores are spread out from the country's average score.

- Seventeen countries, or about half of the countries participating in PIRLS 2001, show less variation in student performance than the United States. Ten countries show a higher variation, while the remaining eight countries show no detectable differences in variation in student performance compared to the United States.

BEST COPY AVAILABLE

Figure 4. Distribution of average combined reading literacy scale scores of fourth-graders by percentiles, by country: 2001



[1] Met guidelines for sample participation rates only after replacement schools were included.

[2] Nearly satisfying guidelines for sample participation rates after replacement schools were included.

[3] National Desired Population does not cover all of International Desired Population because coverage falls below 65 percent.

[4a] National Defined Population covers less than 95 percent of National Desired Population.

[4b] National Defined Population covers less than 80 percent of National Desired Population.

[5] Canada is represented by the provinces of Ontario and Quebec (O, Q) only.

[6] Hong Kong, SAR is a Special Administrative Region (SAR) of the People's Republic of China.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

Reading Literacy by Benchmarks

Average scores in figure 3 indicate how well the United States performs relative to other countries, but the scores do not indicate the proficiency required to reach a particular score. To gain a better understanding of what scores represent in terms of reading proficiency, PIRLS 2001 selected four cutoff points on the combined reading literacy scale labeled *international benchmarks*. These benchmarks were selected to correspond to the score points at or above which the lower quarter, median, upper quarter, and top 10 percent of fourth-graders in the international PIRLS 2001 sample performed.[4]

Student responses at the four benchmarks were analyzed to describe a set of reading skills and strategies displayed by fourth-graders at those points. These descriptions, together with the cut point scores, are listed in figure 5.[5]

- On the combined reading literacy scale, 19 percent of the fourth-grade students in the United States reach the top 10 percent benchmark, 41 percent the upper quarter benchmark, 68 percent the median benchmark, and 89 percent the lower quarter benchmark (figure 6). The percentage of U.S. fourth-graders reaching each of these benchmarks is higher than the international averages.

Figure 5. Fourth-graders' reading skills and strategies, and cut point scores, by benchmark points for the combined reading literacy scale: 2001

Benchmark	Cut Point Scores	Reading skills and strategies*
Top 10 percent	615 and above	• Demonstrate ability to integrate ideas and information • Provide interpretations about characters' feelings and behaviors with text-based support • Integrate ideas across the text to explain the broader significance or theme of the story • Demonstrate understanding of informational materials by integrating information across various types of materials and successfully applying it to real-world situations
Upper quarter	570 and above	• Demonstrate ability to make inferences and recognize some text features in literary texts • Make inferences to describe and contrast characters' actions
Median	510 and above	• Make elementary interpretations • Locate specific parts of text to retrieve information • Make observations about whole texts
Lower quarter	435 and above	• Retrieve explicitly stated details from various literary and informational texts

* The responses of students who score within 5 points of each of the cut point scores were evaluated to determine reading skills and strategies displayed by fourth-graders at those points. Procedures used for anchoring these items to the benchmarks are explained more fully in the *PIRLS Technical Report* at **www.pirls.org**.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

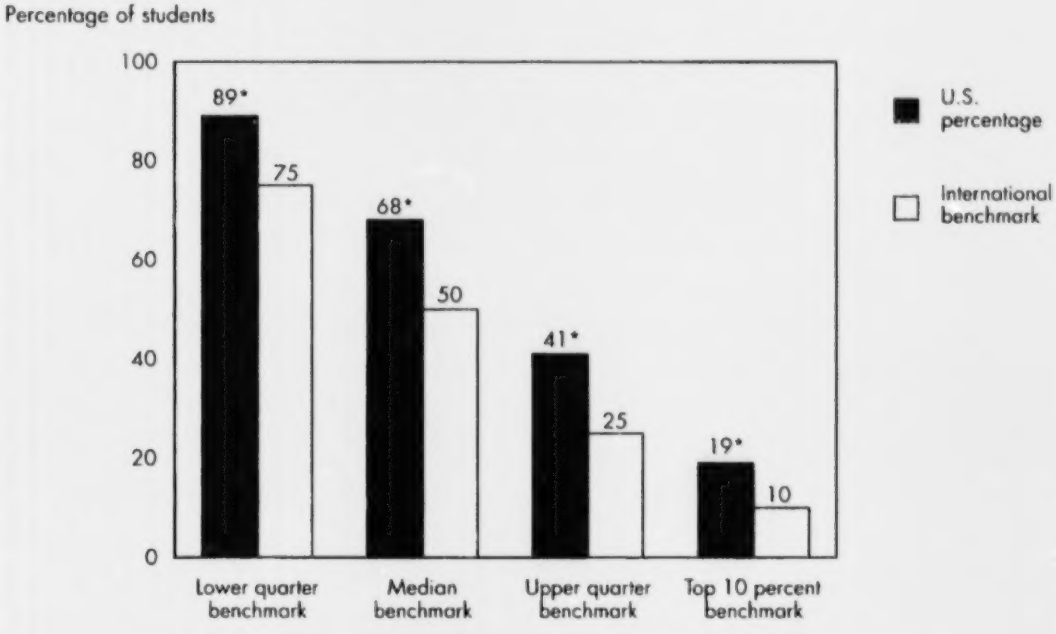
[4] Benchmarking in PIRLS describes the performance of students at four international benchmarks based on the distribution of scores and the pattern of items answered correctly. Proficiency levels for the National Assessment of Educational Progress (NAEP) (i.e., Basic, Proficient, and Advanced) are established by the National Assessment Governing Board based on recommendations from broadly representative panels of educators and the general public who determine what students should know and be able to do at the three levels of performance in each subject area and in each grade assessed.

[5] If students' reading achievement was distributed in the same way in every country, then each country would be expected to have approximately 10 percent of fourth-graders reading the top 10 percent benchmark, 25 percent the upper quarter benchmark, 50 percent the median benchmark, and 75 percent the lower quarter benchmark.

BEST COPY AVAILABLE

- Compared with the United States, no other country but England (24 percent) reports a higher percentage of students at the top 10 percent benchmark on the combined reading literacy scale. Sweden (47 percent) reports a higher share of students at the upper quarter benchmark compared to the United States.
- On the literary subscale, for the United States, 22 percent of the students reach the top 10 percent benchmark, 43 percent the upper quarter benchmark, 70 percent the median benchmark, and 90 percent the lower quarter benchmark. The percentage of U.S. fourth-graders reaching each of these benchmarks on the literary subscale is higher than the corresponding international averages.
- On the informational subscale, for the United States, 15 percent of the students reach the top 10 percent benchmark, 36 percent the upper quarter benchmark, 66 percent the median benchmark, and 89 percent the lower quarter benchmark. The percentage of U.S. fourth-graders reaching these benchmarks on the informational subscale is higher than the corresponding international averages.

Figure 6. Percentage of fourth-grade students who reach the PIRLS 2001 achievement benchmarks: 2001



* Significant difference between U.S. percentage and international benchmark in this category.

NOTE: The United States met guidelines for sample participation rates after replacement schools were included.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

How Different Groups Perform

Achievement by Sex

In the United States and many other countries, policymakers and educators are interested not only in overall achievement but also in achievement by specific groups of students. For example, patterns of differences between boys and girls in reading achievement across countries can point to areas where additional educational resources might be focused.

- Fourth-grade girls score higher than fourth-grade boys on the combined reading literacy scale on average in every participating PIRLS 2001 country (figure 7). In the United States, on average, girls score 18 points higher than boys on the combined reading literacy scale. Internationally, the average score difference between boys and girls range from 8 points (Italy) to 27 points (Belize, Iran, and New Zealand).[6]

- Fourth-grade girls score higher than boys on both the literary and informational subscales in all of the participating PIRLS 2001 countries (figure 8). In the United States, fourth-grade girls, on average, outscore boys by 16 points on both the literary and informational subscales.

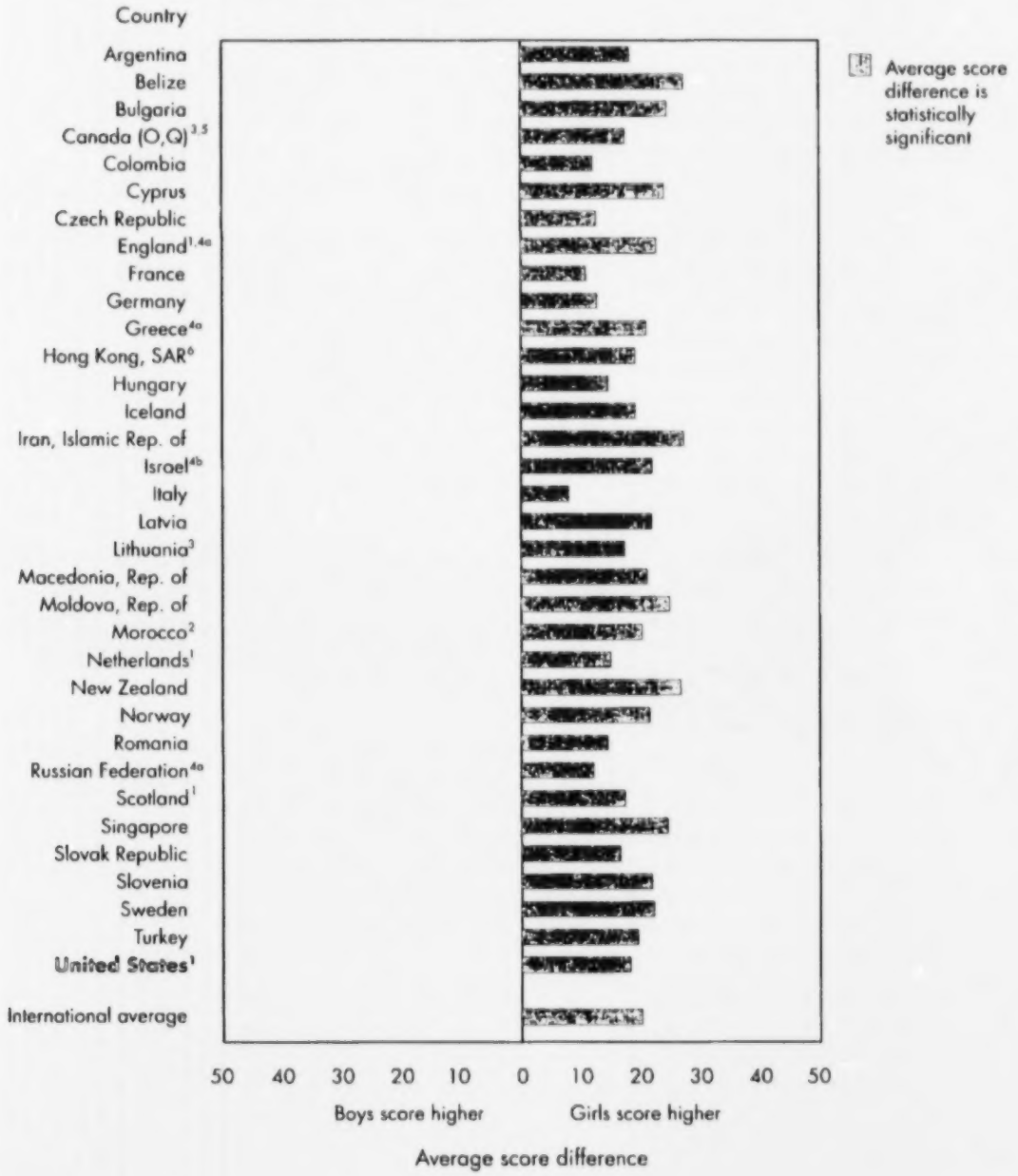
- Fourth-grade girls in Sweden, England, the Netherlands, and Bulgaria outscore U.S. girls on the combined reading literacy scale. However, U.S. girls perform better than their counterparts in 21 of the participating PIRLS 2001 countries.

- Fourth-grade boys in the Netherlands and Sweden outperform U.S. boys on the combined reading literacy scale, although U.S. boys perform better than their peers in 22 of the participating PIRLS 2001 countries.

[6] Differences in scores by sex are not shown here for Kuwait due to low response rates on the question related to sex. However, the international average includes Kuwait's average scale score.

BEST COPY AVAILABLE

Figure 7. Difference in average scores between boys and girls for the combined reading literacy scale of fourth-graders, by country: 2001



[1] Met guidelines for sample participation rates only after replacement schools were included.

[2] Nearly satisfying guidelines for sample participation rates after replacement schools were included.

[3] National Desired Population does not cover all of International Desired Population because coverage falls below 65 percent.

[4a] National Defined Population covers less than 95 percent of National Desired Population.

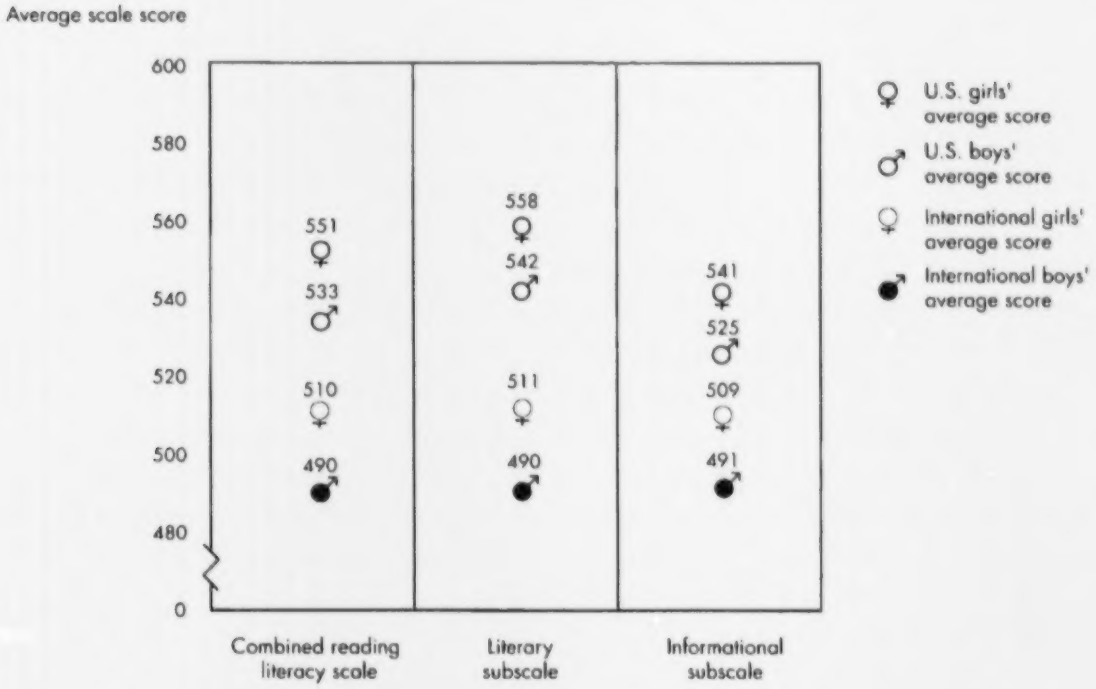
[4b] National Defined Population covers less than 80 percent of National Desired Population.

[5] Canada is represented by the provinces of Ontario and Quebec (O, Q) only.

[6] Hong Kong, SAR is a Special Administrative Region (SAR) of the People's Republic of China.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

Figure 8. Fourth-graders' average scores on the combined reading literacy scale, literary subscale, and informational subscale for boys and girls: 2001



NOTE: The United States met guidelines for sample participation rates after replacement schools were included.

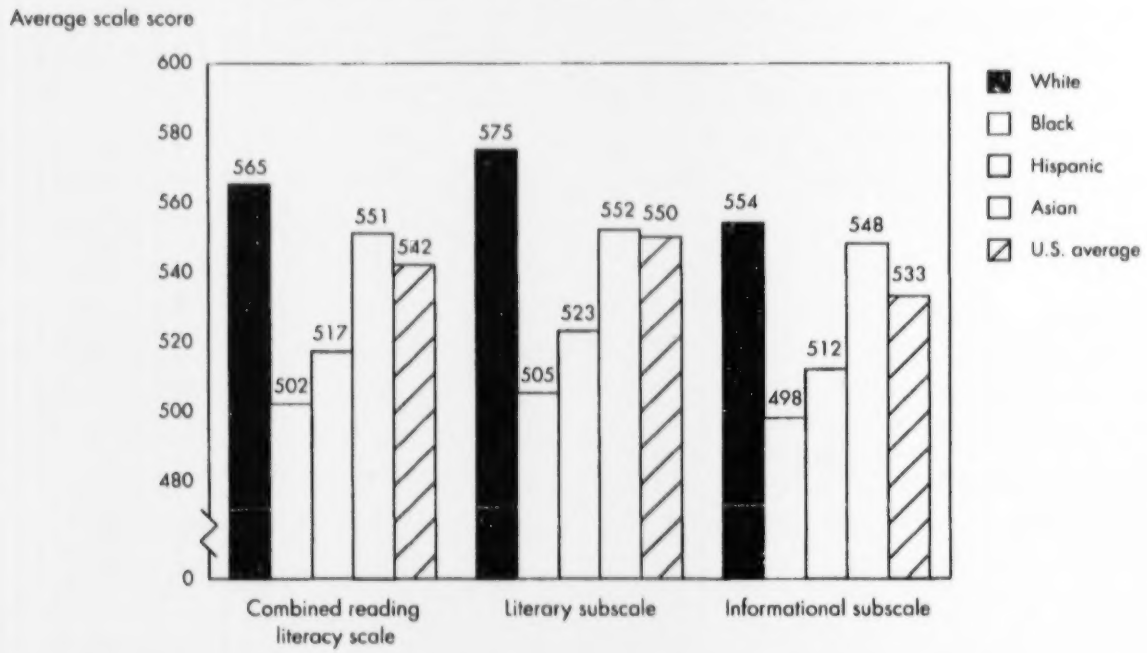
SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

U.S. Achievement by Race/Ethnicity

Another area of interest among policymakers and educators is the achievement of racial/ethnic groups. A number of countries that participated in PIRLS 2001 have large and diverse racial/ethnic groups. However, since these groups vary considerably across countries, it is not possible to compare their performance internationally. Thus, the findings on this page refer only to PIRLS 2001 results for the United States.

- With the exception of Black fourth-graders, each racial/ethnic group in the United States scores higher than the international average (i.e., 500) on the combined reading literacy scale, as well as on the two reading subscales.
- There is considerable variation in scores among the racial/ethnic groups in the United States. On average, White fourth-grade students perform better than Black and Hispanic fourth-graders on the combined reading literacy scale, as well as on the two subscales (figure 9). Asian fourth-grade students, on average, also perform better than Black and Hispanic students on the combined reading literacy scale, as well as on the informational subscale. On the literary subscale, Asian students perform better than Black students, while there are no detectable differences in performance between Asian and Hispanic students. There are no detectable differences in scores between White and Asian fourth-grade students across any of the reading scales.
- A larger percentage of White fourth-graders in the United States reach the top 10 percent benchmark than do Black or Hispanic fourth-graders. For example, 25 percent of White fourth-graders reach the top 10 percent benchmark, while 6 percent of Black and 10 percent of Hispanic fourth-graders reach the same benchmark. There is no detectable difference in the percentage of White and Asian fourth-graders who reach the top 10 percent benchmark, but a larger percentage of Asian fourth-graders reach that benchmark than do Black fourth-graders.
- A larger percentage of both White and Asian fourth-graders in the United States reach the upper quarter benchmark than do Black and Hispanic fourth-graders. For example, 51 percent of White and 46 percent of Asian fourth-graders reach this benchmark, while 19 percent of Black and 27 percent of Hispanic fourth-graders reach that benchmark.

Figure 9. U.S. fourth-graders' average scores for the combined reading literacy scale, literary subscale, and informational subscale, by race/ethnicity: 2001



NOTE: Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic-origin unless specified. The United States met guidelines for sample participation rates after replacement schools were included.

SOURCE: International Association for the Evolution of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

13

U.S. Achievement by Control of School

- On average, fourth-grade students in private schools in the United States score significantly higher than fourth-grade students in public schools on the combined reading literacy scale, and also on the literary and informational subscales. For example, on the combined reading literacy scale and the informational subscale, on average, fourth-grade students in private schools score 42 points higher than students in public schools. On the literary subscale, private schools fourth-graders score an average 45 points higher than public school fourth-graders.

U.S. Achievement by Poverty Level in Public Schools

One measure of poverty in U.S. public elementary schools is the percentage of students eligible for free or reduced-price lunch.[7] In order to examine how fourth-graders' scores on the combined reading literacy scale are associated with their schools' poverty level (percentage of students receiving free or reduced-price lunch), U.S. public schools were classified into five groups: (1) schools with the lowest poverty levels of less than 10 percent; (2) schools with poverty levels ranging from 10 to 24.9 percent; (3) schools with poverty levels ranging from 25 to 49.9 percent; (4) schools with poverty levels ranging from 50 to 74.9 percent; and (5) schools with the highest poverty levels of 75 percent or more. Again, data on this page refer only to PIRLS 2001 results from public schools in the United States.[8]

- Fourth-graders in U.S. public elementary schools with the highest poverty levels score lower on the combined reading literacy scale compared to their counterparts in schools with lower poverty levels (figure 10).

- Fourth-graders in schools with intermediate poverty levels of 10 to 24.9 percent and 25 to 49.9 percent score higher on the combined reading literacy scale than students in schools with poverty levels of 50 to 74.9 percent and 75 percent or more. However, there are no detectable differences in scores between U.S. fourth-graders in public schools with poverty levels of 10 to 24.9 percent and 25 to 49.9 percent.

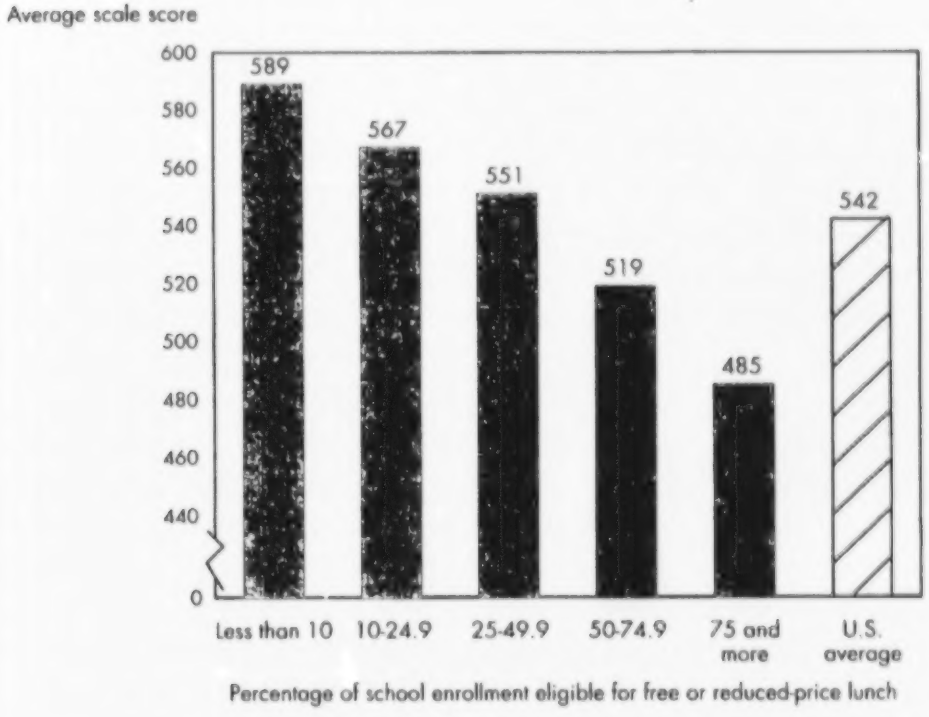
- On average, lower percentages of fourth-graders in the highest poverty public schools in the United States reach the upper two international benchmarks (top 10 percent and upper quartile) than their counterparts in the lowest poverty schools. For example, in the highest poverty schools, about 3 percent of the fourth-grade students reach the top 10 percent international benchmark, while in the lowest poverty schools, about 34 percent of the students reach the same benchmark. Additionally, about 14 percent of students in the highest poverty schools reach the upper quarter benchmark, but in the lowest poverty schools, 64 percent of the students reach that benchmark.

[7] Data for the percentage of students eligible for free or reduced-price lunch in U.S. public elementary schools participating in PIRLS 2001 were taken from the U.S. Department of Education, National Center for Education Statistics (NCES) Common Core of Data (CCD), "Public Elementary/Secondary School Universe Survey," 1999–2000.

[8] Since the measure of school poverty used for the United States in this analysis cannot be applied to other countries, only data for U.S. schools are used in these comparisons.

BEST COPY AVAILABLE

Figure 10. U.S. public school fourth-graders' average scores for the combined reading literacy scale, by the percentage of school enrollment eligible for free or reduced-price lunch: 2001



NOTE: The United States met guidelines for sample participation rates after replacement schools were included.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001, unpublished tables; United States Department of Education, National Center for Education Statistics (NCES) Common Core of Data (CCD), "Public Elementary/Secondary School Universe Survey," 1999–2000.

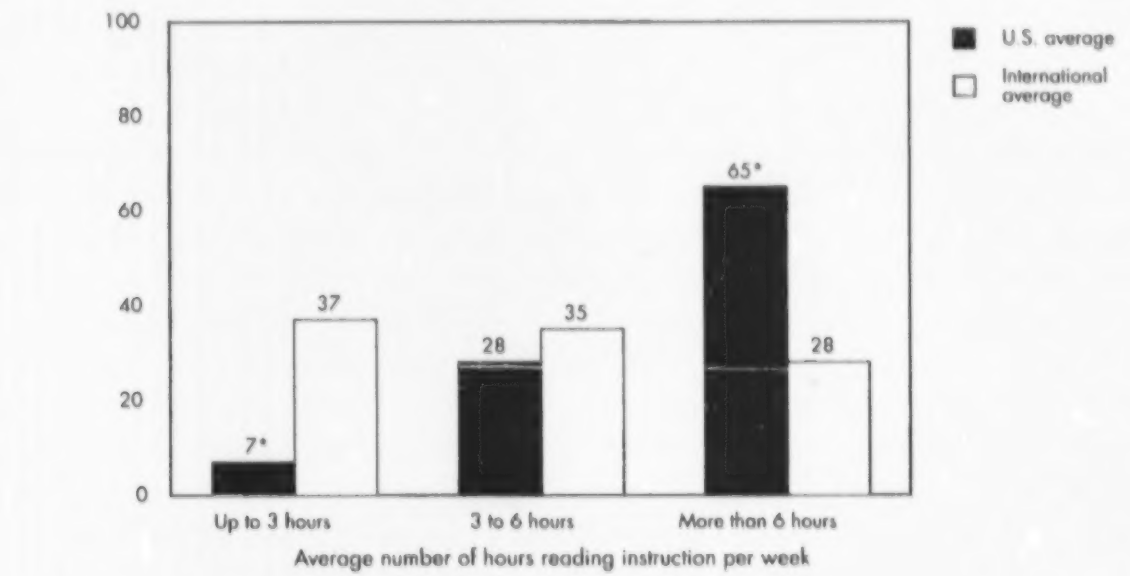
Reading and Instruction in the Classroom

Reading Curriculum and Instructional Time

Do school principals and teachers encourage reading instruction through a variety of initiatives? What proportion of the school day is spent in reading instruction? Answers to these questions can give an indication of the emphasis that reading instruction receives in the curriculum of a country.

- According to school principals, 72 percent of U.S. fourth-graders attend schools that have a written statement describing the reading curriculum, which is nearly double the international average of 37 percent.
- Almost all (95 percent) U.S. fourth-grade students attend schools with a curricular emphasis on reading. This is greater than the international average of 78 percent.
- Principals report that 95 percent of U.S. fourth-grade students attend schools with informal initiatives to encourage reading, which is greater than the international average of 76 percent.[9]
- Based on teacher reporting, 65 percent of U.S. fourth-graders receive more than 6 hours of reading instruction per week, a higher percentage than the international average of 28 percent (figure 11). This percentage is also higher than the national average in 31 of the other 34 participating PIRLS 2001 countries.

Figure 11. Percentage of fourth-graders by average number of hours of reading instruction each week: 2001



* Significant difference between U.S. average and international average in this category.

NOTE: The United States met guidelines for sample participation rates after replacement schools were included.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

[9] Informal initiatives to promote reading include book clubs, independent reading contests, and schoolwide recreational reading periods to encourage students to read.

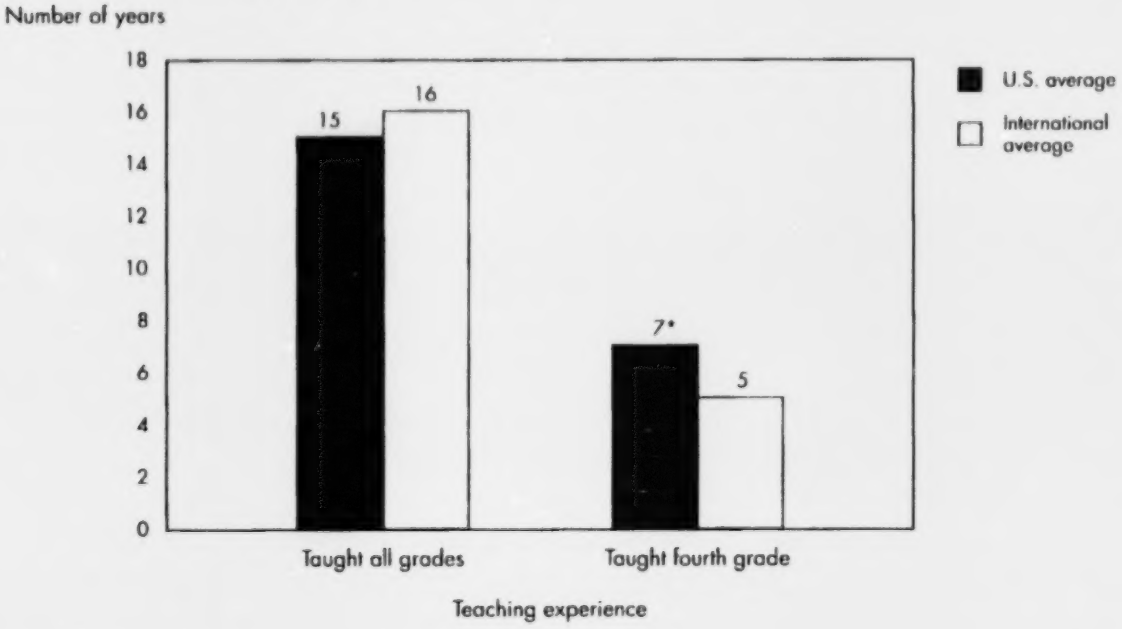
- The average combined reading literacy achievement scores of U.S. fourth-graders do not vary by the amount of instructional time they receive.

Teacher Preparation and Experience

Examining teachers' preparation and tenure indicates the experience of teachers in the classroom. On the teacher questionnaire in PIRLS 2001, teachers were asked about the training they have received and the number of years they have been teaching.

- Based on teacher reports of their preparation for teaching, 95 percent of U.S. fourth-graders are taught by certified teachers.[10] This is higher than the corresponding international average of 89 percent.
- U.S. fourth-graders appear to be taught by teachers who have more experience teaching fourth grade than their counterparts in the majority of the participating PIRLS 2001 countries. On average, U.S. fourth-grade students are taught by teachers who have been teaching fourth grade for 7 years (figure 12).[11] Twenty-six of the other 34 participating countries reported that their fourth-graders are taught by teachers with fewer years of experience teaching fourth grade.

Figure 12. Average number of years fourth-grade students' teachers have taught all grades and fourth grade: 2001



* Significant difference between U.S. average and international average in this category.

NOTE: The United States met guidelines for sample participation rates after replacement schools were included.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

[10] Indicates that students are taught by a teacher with a teaching certificate. The most recent NAEP reading assessment data from 1994 show that 95 percent of the teachers of fourth-grade students were certified in the state in which they taught. In the 2001 Schools and Staffing Survey (SASS), 97 percent of fourth-grade teachers reported that they were certified.

[11] In the 2001 SASS, fourth-grade teachers reported that on average they had been teaching for 14 years.

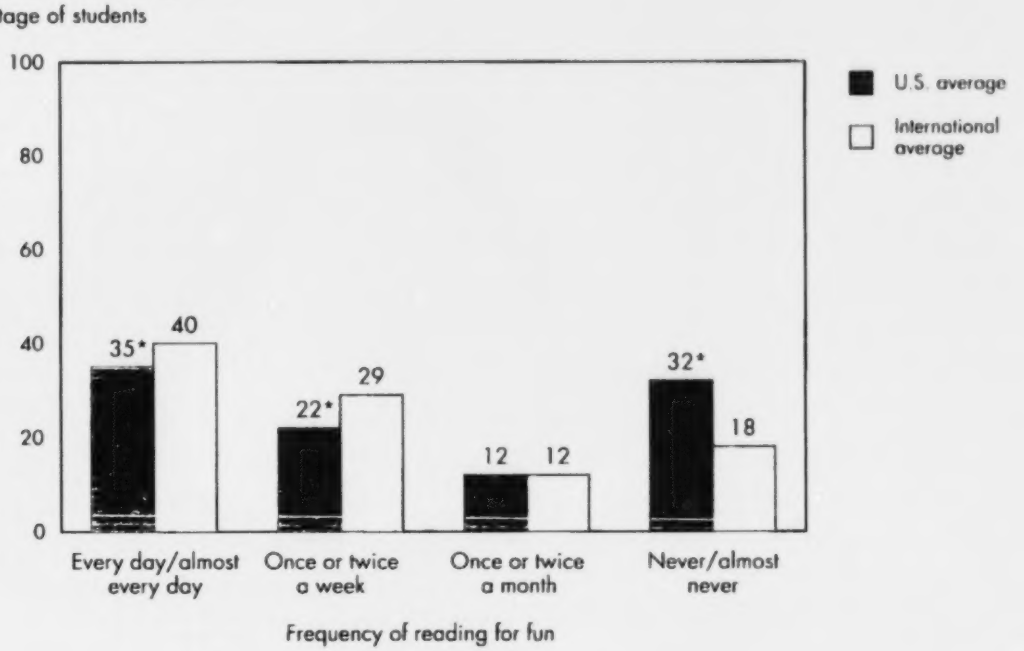
Reading Outside of School

Reading Outside of School for Enjoyment

To investigate the reading habits of fourth-graders outside of school, PIRLS asked students a series of questions about whether they read for fun outside of school and how often they did so. Students could indicate that they read for fun "every day or almost every day," "once or twice a week," "once or twice a month," or "never or almost never."

- Thirty-five percent of U.S. fourth-graders report reading for fun every day or almost every day. This percentage is smaller than the international average of 40 percent.
- Thirty-two percent of U.S. fourth-graders report that they never or almost never read for fun outside of school, a significantly higher percentage than the international average of 18 percent (figure 13).
- In the United States, fourth-graders who read for fun every day or almost every day have higher average scores on the combined reading literacy scale compared to those who never or almost never read for fun, or do so once or twice a month. This pattern holds at the international level as well, based on the international averages.

Figure 13. Percentage of fourth-grade students who read for fun outside of school, by frequency of reading: 2001



* Significant difference between U.S. average and international average in this category.

NOTE: Detail may not sum to totals due to rounding. The United States met guidelines for sample participation rates after replacement schools were included.

SOURCE: International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS) 2001.

Choice of Activities Outside of School

To learn more about students' reading habits, PIRLS 2001 asked students about their choice of reading materials and how often they read different types of texts when they are not in school.

- In the United States, 92 percent of fourth-graders report reading for information at least once or twice a month, a higher percentage than those who report reading either literary fiction, such as stories or novels (79 percent), or comics (43 percent) at least once or twice a month.
- In the United States, 43 percent of fourth-graders report that they read comics at least once or twice a month, a significantly lower percentage than the international average of 74 percent.
- U.S. fourth-graders who report reading literary fiction outside of school at least once or twice a month have higher scores on the combined reading literacy scale than those who never or almost never do so. This pattern is also evident at the international level, based on international averages.
- No measurable differences in scores on the combined reading literacy scale are detected between U.S. fourth-graders who read informational materials every day or almost every day, and those who never or almost never do so.

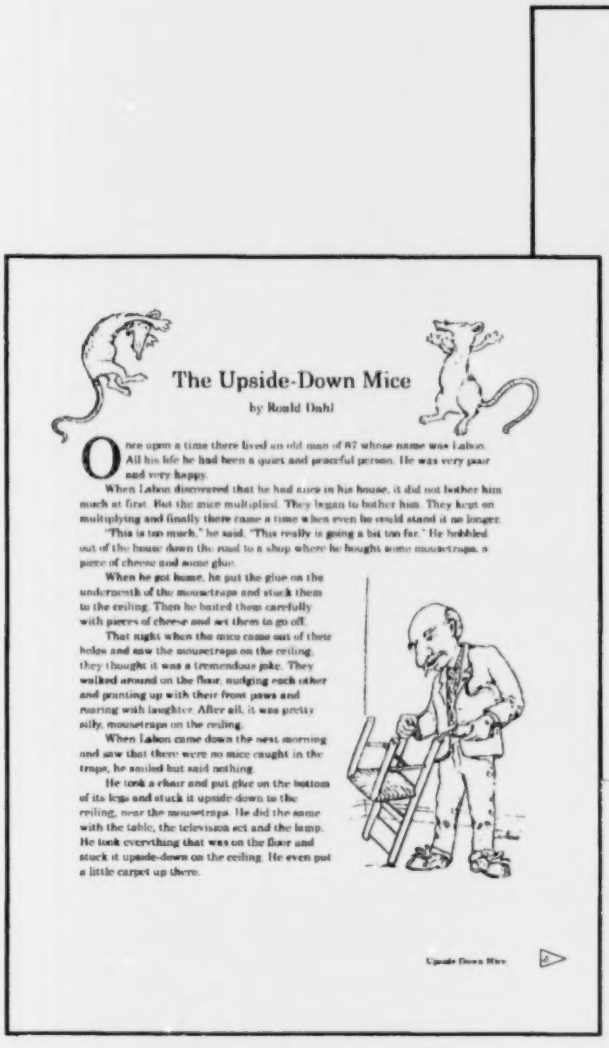
PIRLS 2001 also asked students about their TV- and video-watching habits.

- Eighteen percent of U.S. fourth-graders report watching TV or videos on a normal school day for 5 hours or more. This is significantly higher than the international average of 12 percent. On average, U.S. fourth-graders report watching TV or videos daily for a greater number of hours than the international average (2.2 hours vs. 2 hours, respectively).
- Looking at the international average for the combined reading literacy scale, fourth-graders who watch TV for more than 5 hours on a normal school day score lower than those who watch TV for 3 to 5 hours a day or less frequently. In the United States, the same finding holds.

BEST COPY AVAILABLE

Sample Items from PIRLS 2001

Exhibits 4 through 9 contain reading passages from stories in PIRLS 2001 and a number of assessment items (see below). The items show actual student responses and compare U.S. fourth-graders' performance to the international average. The items also demonstrate acceptable performance at the four benchmarks (top 10 percent, upper quarter, median, and lower quarter). The reading passage and all of these items have been released to the public by IEA.



The Upside-Down Mice

by Roald Dahl

Once upon a time there lived an old man of 87 whose name was Labon. All his life he had been a quiet and peaceful person. He was very poor and very happy.

When Labon discovered that he had mice in his house, it did not bother him much at first. But the mice multiplied. They began to bother him. They kept on multiplying and finally there came a time when even he could stand it no longer.

"This is too much," he said. "This really is going a bit too far." He hobbled out of the house down the road to a shop where he bought some mousetraps, a piece of cheese and some glue.

When he got home, he put the glue on the underneath of the mousetraps and stuck them to the ceiling. Then he baited them carefully with pieces of cheese and set them to go off.

That night when the mice came out of their holes and saw the mousetraps on the ceiling, they thought it was a tremendous joke. They walked around on the floor, nudging each other and pointing up with their front paws and roaring with laughter. After all, it was pretty silly, mousetraps on the ceiling.

When Labon came down the next morning and saw that there were no mice caught in the traps, he smiled but said nothing.

He took a chair and put glue on the bottom of its legs and stuck it upside-down to the ceiling, near the mousetraps. He did the same with the table, the television set and the lamp. He took everything that was on the floor and stuck it upside-down on the ceiling. He even put a little carpet up there.

Upside Down Mice

The next night when the mice came out of their holes they were still joking and laughing about what they had seen the night before. But now, when they looked up at the ceiling, they stopped laughing very suddenly.

"Good gracious me!" cried one. "Look up there! There's the floor!"

"Heavens above!" shouted another. "We must be standing on the ceiling!"

"I'm beginning to feel a little giddy," said another.

"All the blood's going to my head," said another.

"This is terrible!" said a very senior mouse with long whiskers. "This is really terrible! We must do something about it at once!"

"I shall faint if I have to stand on my head any longer!" shouted a young mouse.

"Me too!"

"I can't stand it!"

"Save us! Do something, somebody, quick!"

They were getting hysterical now. "I know what we'll do," said the very senior mouse. "We'll all stand on our heads, then we'll be the right way up."

Obediently, they all stood on their heads, and after a long time, one by one they fainted from a rush of blood to their brains.

When Labon came down the next morning the floor was littered with mice. Quickly he gathered them up and popped them all in a basket.

So the thing to remember is this: whenever the world seems to be terribly upside-down, make sure you keep your feet firmly on the ground.

Upside Down Mice

BEST COPY AVAILABLE

Exhibit 4. Example of item at the top 10 percent PIRLS 2001 international benchmark

2 out of 3 Points: Partial Credit Sample Response and Results

12. You learn what Labon is like from the things he does. Describe what he is like and give **two** examples of what he does that show this.

He was smart. He thought of a good way to trick the mice.

	Average percent correct
U.S. average	49*
International average	30

* Significant difference between U.S. average and international average.

Exhibit 5. Example of item at the top 10 percent PIRLS 2001 international benchmark

1 Point: Full Credit Sample Response and Results

12. Why did Labon smile when he saw there were no mice in the traps?

Labon knew the mice did not know his trick yet.

	Average percent correct
U.S. average	47*
International average	31

* Significant difference between U.S. average and international average.

BEST COPY AVAILABLE

Exhibit 6. Example of item at the upper quarter PIRLS 2001 international benchmark

1 Point: Full Credit Sample Response and Results

11. Do you think the mice were easy to fool? Give **one** reason why or why not.

1 No It took two nights to trick them.

	Average percent correct
U.S. average	54*
International average	37

* Significant difference between U.S. average and international average.

Exhibit 7. Example of item at the median PIRLS 2001 international benchmark

1 Point: Full Credit Sample Response and Results

13. Which words best describe this story?

- ○ serious and sad
- ○ scary and exciting
- ● funny and clever
- ○ thrilling and mysterious

	Average percent correct
U.S. average	81*
International average	68

* Significant difference between U.S. average and international average.

BEST COPY AVAILABLE

Exhibit 8. Example of item at the lower quarter PIRLS 2001 international benchmark

1 Point: Full Credit Sample Response and Results

12. Where did Labon put the mice when he picked them up from the floor?

(1) In a basket

	Average percent correct
U.S. average	87
International average	84

Exhibit 9. Example of item at the lower quarter PIRLS 2001 international benchmark

1 Point: Full Credit Sample Response and Results

1. Why did Labon want to get rid of the mice?

- ○ He had always hated mice.
- ● There were too many of them.
- ○ They laughed too loudly.
- ○ They ate all his cheese.

	Average percent correct
U.S. average	84*
International average	79

* Significant difference between U.S. average and international average.

BEST COPY AVAILABLE

IEA International Reading Literacy Study of 1991

Reading Performance Over Time

Ten years before PIRLS 2001 was administered, the IEA conducted the IEA International Reading Literacy Study of 1991. This study, like PIRLS 2001, assessed the reading literacy of fourth-graders in over 30 countries using 42 items taken from 6 reading passages. However, when a followup for the 1991 study was being planned, the IEA decided to discontinue it and develop a new assessment incorporating the latest approaches to measuring reading literacy (Campbell et al., 2001). This new study would become PIRLS 2001.

While participating in PIRLS 2001, some countries expressed interest in comparing reading performance between 1991 and 2001. Since comparisons between the two assessments were impossible (see exhibit 10), the IEA gave participating countries an opportunity to readminister the 1991 study during the PIRLS 2001 administration. This readministered study was identical in content, timing, and directions to that given to students in 1991 and allowed comparisons of the performance of students in 2001 with those in 1991. A separate sample of students was drawn in each country so as not to overburden students assessed in PIRLS 2001. Nine countries, including the United States, participated in what is called the IEA International Reading Literacy Study of 1991 (figure 14).

Exhibit 10. Interpreting PIRLS 2001 in light of the IEA International Reading Literacy Study of 1991

In anticipation of the simultaneous release of PIRLS 2001 and the IEA International Reading Literacy Study of 1991, NCES commissioned a comparative analysis of the two assessments. Frameworks, passages, and items in both studies were reviewed and compared. Results indicate that the two studies are quite different. To cite a few examples: Reading passages in PIRLS 2001 were found to be "longer, more engaging, and more complex in most cases" than those found in the IEA International Reading Literacy Study of 1991 (Kapinus, 2003, p. 8). PIRLS 2001 also used many more constructed-response (essay-type) questions and presented them in a way "that might have improved students' motivation to read and respond to the texts" (Kapinus, 2003, p. 8). The analysis also found that, in general, PIRLS 2001 tapped skills "requiring deeper thinking" than those in the IEA International Reading Literacy Study of 1991 (Kapinus, 2003, p. 8). Because of these and other differences, it is impossible to directly compare results from these two assessments. However, separately, each study provides important clues about how well students in these countries, including U.S. fourth-graders, perform in reading literacy.

Figure 14. Participating countries in the IEA International Reading Literacy Study of 1991: 2001

Greece	Italy	Slovenia
Hungary	New Zealand	Sweden
Iceland	Singapore	**United States**

SOURCE: International Association for the Evaluation of Educational Achievement, IEA International Reading Literacy Study of 1991.

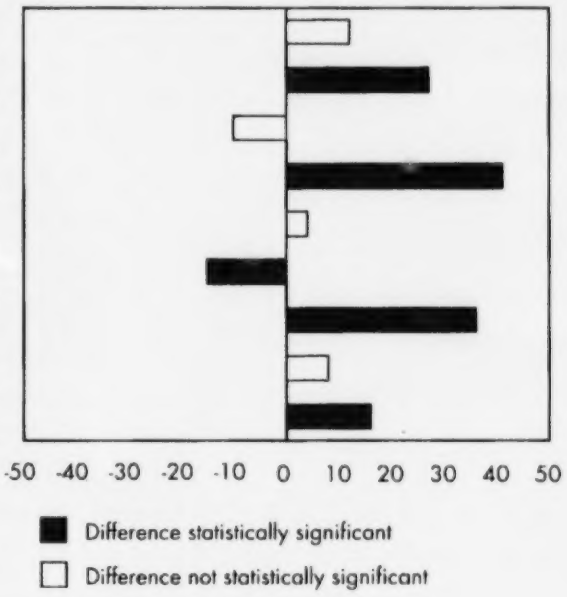
BEST COPY AVAILABLE

Performance on the IEA International Reading Literacy Study of 1991

- Based on the readministration of the 1991 study in 2001, no detectable change is observed in the achievement of fourth-graders on the combined reading literacy scale in the United States in 2001 compared to 1991.
- Fourth-graders in five of the nine participating countries perform significantly better, on average, on the 1991 study combined reading literacy scale in 2001 compared to 1991, while fourth-graders in three countries show no detectable difference in average achievement between 1991 and 2001 (figure 15). One country, Sweden, has a significantly lower average score in 2001 than in 1991.

Figure 15. Average scores for the IEA International Reading Literacy Study of 1991 combined reading literacy scale and differences in average achievement scores for fourth-graders, by country: 1991 and 2001

Country	2001 average scale score	1991 average scale score	1991-2001 difference
Italy	513	500	12
Iceland	513	486	27
United States	**511**	**521**	**-10**
Greece	507	466	41
New Zealand	502	498	4
Sweden	498	513	-15
Slovenia	493	458	36
Singapore	489	481	8
Hungary	475	459	16


Difference in average achievement between 1991 and 2001
-50 -40 -30 -20 -10 0 10 20 30 40 50
Difference statistically significant
Difference not statistically significant

NOTE: Countries are ordered based on the 2001 average score.

SOURCE: International Association for the Evaluation of Educational Achievement, IEA International Reading Literacy Study of 1991, 2001.

Works Cited

Binkley, M., and Kelly, D.L. (forthcoming). *A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments*. U.S. Department of Education. Washington, DC: NCES Working Paper.

Campbell, J.R., Kelly, D.L., Mullis, I.V.S., Martin, M.O., and Sainsbury, M. (2001). *Framework and Specifications for PIRLS Assessment 2001*. Chestnut Hill, MA: Boston College.

Kapinus, B. (2003). *PIRLS-IEA Reading Literacy Framework: Comparative Analysis of the 1991 IEA Reading Study and PIRLS* (NCES 2003—005). U.S. Department of Education. Washington, DC: NCES Working Paper.

Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., and Kennedy, A.M. (2003). *Trends in Children's Literacy Achievement 1991–2001: IEA's Repeat in Nine Countries of the 1991 Reading Literacy Study*. Chestnut Hill, MA: Boston College.

Martin, M.O., Mullis, I.V.S., and Kennedy, A.M. (Eds.) (2003). *PIRLS 2001 Technical Report*. Chestnut Hill, MA: Boston College.

Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., and Kennedy, A.M. (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools in 35 Countries*. Chestnut Hill, MA: Boston College.

Appendix A

Technical Notes

Most of these notes are derived from the *PIRLS Technical Report*. For a full discussion of these topics, see the *PIRLS Technical Report* at **www.pirls.org**.

Background

PIRLS 2001 formed the Reading Development Group (RDG) to help construct the PIRLS 2001 Framework and to review and endorse the final reading test. The RDG consisted of Marilyn Binkley, Karl Blueml, Sue Horner, Pirjo Linnakylo, Martine Remond, Keen See Tan, and William Tunmer provided valuable support in the design of the assessment. Jay Campbell of Educational Testing Service served as a technical advisor to the RDG.

The RDG concluded that at least 4 hours of assessment material and 2 hours for each reading purpose (literary and informational) were needed to provide a valid and reliable measure of reading achievement. Since it would not be possible to administer the entire test to any one child, PIRLS 2001 used a matrix sampling technique to distribute the assessment material among students, yet retain linkages necessary for scaling the achievement data.

Assessment Design

The reading material was divided into 40-minute "blocks," each comprising a story or article and items representing at least 15 score points. There were eight such blocks, four for each reading purpose: literary and informational. The eight assessment blocks were distributed across 10 test booklets, and each student completed one booklet in an 80-minute testing session. Each booklet contained two blocks.

One of the 10 booklets was the PIRLS 2001 Reader, a color booklet containing two reading passages, which appeared only in that booklet. The distribution of blocks across booklets linked the booklets to enable the achievement data to be scaled using item response theory methods.

Student Population Assessed

In 30 of the 35 PIRLS 2001 countries, including the United States, the students who completed the assessment had received 4 years of formal schooling. Five countries that assessed students who had received formal schooling for a different number of years were Belize (4 or 5 years), England (5 years), the Russian Federation (3 or 4 years), Scotland (5 years), and Slovenia (3 years). However, all participating PIRLS 2001 countries assessed the reading literacy of their students in the upper of two grades with the most 9-year-olds.

In the United States, the mean age of the students who completed the PIRLS 2001 assessment was 10.2 years. The average mean age of students in countries with significantly lower average scores than the United States on the combined reading literacy scale was also 10.2 years, and the mean age of participating students in countries with significantly higher average scores than the United States was 10.4 years.

Passages

The reading passages formed the foundation of the reading literacy test. In accordance with the framework, four assessment blocks contained literary texts and four contained informational texts. The passages were authentic texts drawn from children's storybook and informational sources. Submitted and reviewed by PIRLS 2001 participating countries, the literary passages included realistic stories and traditional tales. The informational texts included chronological and nonchronological articles, a biographical article, and an informational leaflet.

PIRLS 2001 examined all passages and test items for cultural bias. A large number of passages were initially provided by participating countries. Only those that were selected by all countries became a part of PIRLS 2001. Test items were examined for individual item statistics and item-by-country interactions. The analysis indicated that items were considered unbiased to a similar extent by all of the participating countries.

Item Development and Scoring

Two item formats were used to assess children's reading literacy: multiple choice and constructed response. Each type of item was used to assess both reading purposes and all four reading processes. Multiple-choice items provided students with four possible answers, one of which was correct. Each multiple-choice item was worth one point. Constructed-response items required students to construct their answers rather than select from among possible answers. These items were worth one, two, or three points, depending on the depth of understanding or extent of textual support the item required.

Each block of assessment material contained from 11 to 14 items that together represented at least 15 score points. Altogether, the PIRLS 2001 reading test included 98 items representing 119 score points–enough to estimate achievement reliably.

Scale anchoring is a way of describing students' performance at different points on the reading achievement scale in terms of the types of texts they were asked to read and the types of items they answered successfully. It involved an empirical component in which items that discriminate between successive points on the scale were identified, and a judgmental component in which reading experts examined the content of the texts and items and generalized to students' comprehension skills and strategies. Criteria were applied to the reading achievement scale results to identify the sets of items that students reaching each international benchmark were likely to answer correctly and that those at the next lower benchmark were unlikely to answer correctly.

Translation

The PIRLS 2001 reading test and questionnaires were prepared in English and translated into 31 other languages. Countries were responsible for translating the instruments into their local language or languages following internationally prescribed procedures. To ensure standardization of instruments across countries, PIRLS 2001 undertook an extensive verification process, whereby each country's data collection instrument was independently reviewed and verified by an external translation company engaged by the IEA. Instruments were verified twice, once before the field test and again before the main data collection. Also, statistical analyses of item data were conducted to check for any evidence of differences in student performance across countries that could indicate translation problems. More information about translation issues in the assessment is available in the *PIRLS Technical Report* at **www.pirls.org.**

Sampling

PIRLS 2001 used a three-stage stratified cluster sample design:

1. The first stage was a sample of primary sampling units (geographic units referred to as PSUs).

2. The second stage consisted of a sample of at least 150 schools using probability-proportional-to-size sampling. Schools were stratified by geographical characteristics (such as states or provinces), school type (such as public or private), and the level of urbanization (such as rural or urban). The United States selected 174 schools after substitution for non-responding schools and tested 3,763 fourth-grade students.

3. The third stage consisted of sampling of one or more classrooms from the target grade in sampled schools. The target grade in each country was the upper of the two grades with the most 9-year-olds. In the United States and the majority of other countries, the target grade was the fourth grade. Each fourth-grade classroom in all selected schools had an equal likelihood of being selected. This resulted in a sample size of at least 3,750 students in each country. For more information about the grade levels that were assessed in each country, see the *PIRLS Technical Report* at **www.pirls.org**.

Exclusions in the PIRLS Sample

A major objective of PIRLS was that the target population, the population actually sampled by PIRLS, be as close as possible to the international desired population. Consequently, each country had to account for any exclusion of eligible students from the international desired population. This applied to school-level exclusions as well as within-school exclusions. Within-school exclusions included the following three groups:

Educable mentally disabled students. These are students who were considered, in the professional opinion of the school principal or other qualified staff members, to be educable mentally disabled, or who had been so diagnosed in psychological tests. This

category included students who were emotionally or mentally unable to follow even the general instructions of the PIRLS test. It did not include students who merely exhibited poor academic performance or discipline problems.

Functionally disabled students. These are students who were permanently physically disabled in such a way that they could not perform in the PIRLS tests. Functionally disabled students who could perform were included in the testing.

Non-native-language speakers. These are students who could not read or speak the language of the test and so could not overcome the language barrier of testing. Typically, a student who had received less than 1 year of instruction in the language of the test was excluded, but this definition was adapted in different countries.

School-level exclusions consisted of students in special education schools, students in vocational/technical schools, and students in alternative schools.

The United States produced a within-school exclusion rate of 4.7 percent and a school level exclusion rate of 0.6 percent, for a combined exclusion rate of 5.3 percent. Internationally, combined exclusion rates ranged from no exclusions in Kuwait to 22.4 percent in Israel, with an average rate of 3.8 percent.

Data Collection

Each country was responsible for carrying out all aspects of the data collection, using standardized procedures developed for the study by IEA. Manuals provided explicit instructions on all aspects of the data collection, from contacting sampled schools to packing and shipping materials to the IEA Data Processing Center in Hamburg, Germany. Manuals were also prepared for test administrators and for cooperating school officials. In all participating PIRLS 2001 countries, data were collected in the final months of the 2000–01 school year.

Quality Control

PIRLS 2001 also implemented an international program of site visits, whereby international quality control monitors visited a sample of 15 schools in each country and observed test administrations. PIRLS 2001 National Research Coordinators in each country were also expected to organize national quality control programs based on the international model, to ensure that data across countries were comparable. The national quality control monitors visited random samples of 10 percent of the schools (in addition to those visited by the international quality control monitors) and monitored testing sessions, recording their observations for later analysis.

Statistical Comparisons in This Report

Comparisons made in this highlights report have been tested for statistical significance. For example, in the commonly made comparison of country averages against the average of the United States, tests of statistical significance were used to establish whether or not the observed differences from the U.S. average were statistically significant.

In almost all instances the tests used were standard *t*-tests. These fell into two categories according to the nature of the comparison being made. In simple comparisons of country averages against the U.S. average or against the international average, the following formula was used to compute the *t* statistic:

$$t = (Est_1 - Est_2) / SQRT[(se_1)^2 + (se_2)^2]$$

Est_1 and Est_2 are the estimates being compared (e.g., average of country A and the U.S. average) and se_1 and se_2 are the corresponding standard errors of these averages.

In several places, between-country comparisons of group differences within countries were made. Comparisons of sex differences in other PIRLS 2001 countries against sex differences in the United States is an example. In these instances the following formula was used:

$$t = \frac{[(Est_{11} - Est_{21}) - (Est_{12} - Est_{22})]}{SQRT[(se_{11}^2 + se_{21}^2) + (se_{12}^2 + se_{22}^2)]}$$

Est_{11} and Est_{21} are the estimates being compared within country A (e.g., girls' reading average and boys' reading average), Est_{12} and Est_{22} are the corresponding estimates for the United States, and se_{11}, se_{21}, se_{12}, and se_{22} are their corresponding standard errors.

Appendix B

Examining the Similarities and Differences Between PIRLS and NAEP

Shortly after the release of this report, the National Center for Education Statistics (NCES) will be releasing results for the 2002 National Assessment of Educational Progress (NAEP) fourth-grade reading assessment. In anticipation of expected questions about these two studies, NCES convened an expert panel to compare various aspects of the content of the PIRLS and NAEP assessments. This involved a close examination of how PIRLS and NAEP define reading, the texts used as the basis for the assessments, and the reading processes required of students in each. The similarities and differences between the two are described here.

Similarities

- PIRLS and NAEP define "reading" similarly, as a constructive process.
- PIRLS and NAEP assess reading for a literary experience and reading to be informed.
- PIRLS and NAEP call for students to develop interpretations, make connections across text, and evaluate aspects of what they have read.
- PIRLS and NAEP use literary passages drawn from children's storybooks and informational texts as the basis for the reading assessment.
- PIRLS and NAEP use multiple-choice and constructed-response questions with similar distributions of these types of questions.

	PIRLS	NAEP
Multiple Choice:	47 percent	45 percent
Short Constructed Response:	44 percent	45 percent
Extended Constructed Response:	8 percent	10 percent

Differences

- PIRLS reading passages are, on average, about half the length of the NAEP reading passages: 547 words vs. 1,000 words.
- Results from the Fry Readability Analysis suggest that the PIRLS reading passages are easier than the NAEP passages (grade 5.0 in PIRLS vs. grade 6.9 in NAEP).
- According to Lexile score analysis, the PIRLS passages were determined to be appropriate for the third to fourth grade, and the NAEP passages were determined to be appropriate for the fourth to fifth grade.
- PIRLS calls for more text-based interpretation than NAEP. NAEP places more emphasis on having students take what they have read and connect to other readings or knowledge and to critically evaluate what they have read.

U.S. Department of Education
ED Pubs
8242-B Sandy Court
Jessup, MD 20794-1398

Official Business
Penalty for Private Use, $300

U.S. POSTAGE PAID
U.S. DEPARTMENT
OF EDUCATION
PERMIT NO. G-17

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☒ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").